

The Amino-acid Sequence in the Phenylalanyl Chain of Insulin

1. THE IDENTIFICATION OF LOWER PEPTIDES FROM PARTIAL HYDROLYSATES

By F. SANGER (Beit Memorial Fellow) AND H. TUPPY*

Biochemical Laboratory, University of Cambridge

(Received 17 January 1951)

When insulin is oxidized with performic acid, the —S—S— bridges of the cystine residues are broken by conversion to —SO₃H groups (Sanger, 1949*a*) and the molecule is split into its separate polypeptide chains. From the oxidized insulin two fractions could be isolated: an acidic fraction *A*, which contained only glycyl *N*-terminal residues (see below) and no basic amino-acids, and a basic fraction *B*, having phenylalanyl *N*-terminal residues. These appeared to be the only significant fractions present.

From a study of the partial hydrolysis products of the dinitrophenyl (DNP) derivatives of the two fractions it was possible to determine the sequence of the amino-acids adjoining the *N*-terminal residues and adjoining the lysine residues (Sanger, 1949*b*). In the case of fraction *B* the terminal sequence was shown to be Phe.Val.Asp.Glu and the lysine residues were present in the sequence Thr.Pro.-Lys.Ala (abbreviations for amino-acids are given in Table 1 below). No DNP-peptides were present which did not fit into these sequences, and from an estimation of the yields of the DNP-peptides produced on partial hydrolysis of DNP-insulin it was concluded that all the *N*-terminal phenylalanyl residues of insulin and all the lysine residues were present in the above two sequences respectively, and hence that there was only one type of phenylalanyl polypeptide chain in insulin. Similar though rather less clear-cut results were obtained for fraction *A*. Assuming a molecular weight of 12,000, it was concluded from these experiments that insulin is built up of two identical phenylalanyl polypeptide chains and two identical glycyl chains, these four chains being joined together by six —S—S— bridges. Various possible structures for the molecule were suggested (Sanger, 1949*c*).

The results with the terminal peptides gave a preliminary indication that the fractions *A* and *B* are both relatively homogeneous preparations of molecules having a single polypeptide chain and containing approximately 20 and 30 amino-acid residues respectively. It thus seemed that an investigation of the smaller peptides produced on

partial hydrolysis might yield considerable information about the overall amino-acid sequence in these fractions. Consden, Gordon & Martin (1947) have described a method for the fractionation of lower peptides using paper chromatography which was successfully used to determine the pentapeptide sequence of 'gramicidin S' (Consden, Gordon, Martin & Syngé, 1947). The present paper describes the application of this technique to partial acid hydrolysates of fraction *B* and the determination of a number of amino-acid sequences.

Throughout this paper the abbreviations for the amino-acid residues suggested by Brand & Edsall (1947) are used. These are listed in Table 1. In

Table 1. *Abbreviations for amino-acid residues*

Amino-acid	Abbreviation
Cysteic acid	CySO ₃ H
Aspartic acid	Asp
Glutamic acid	Glu
Serine	Ser
Glycine	Gly
Threonine	Thr
Alanine	Ala
Tyrosine	Tyr
Valine	Val
Leucine	Leu
Phenylalanine	Phe
Proline	Pro
Histidine	His
Lysine	Lys
Arginine	Arg
Ornithine	Orn

referring to peptides of known structure, the abbreviations for the residues are joined by a full stop (e.g. glycylalanine is written Gly.Ala). When two or more residues are included in square brackets the order is unknown. Thus, for instance, Gly.-[Ala, Leu] refers to a peptide or peptides containing glycine, alanine and leucine, in which the free amino group is on the glycine residue, but the relative order of the alanine and leucine residues is unknown. The residues having the free α -amino groups in peptides will be referred to as *N*-terminal residues, those with free α -carboxyl groups as *C*-terminal residues.

* Present address: II. Chemisches Universitätslaboratorium, Vienna, Austria.

METHODS

Fraction *B* was prepared from crystalline ox insulin as previously described (Sanger, 1949*a*).

Hydrolysis and preliminary group separations

Since the partial hydrolysates were too complex for direct fractionation by paper chromatography, it was necessary to carry out certain preliminary group separations to divide them into a number of simpler fractions containing fewer peptides. For this purpose it is advantageous to use methods depending on physical properties of the peptides different from those responsible for the separations on paper chromatograms. We have used adsorption on charcoal, separation on synthetic ion-exchange resins and ionophoretic methods using three- or four-compartment cells. Ionophoresis in silica jelly, which has been successfully used by Consden, Gordon & Martin (1949) to fractionate acidic peptides, was not used since difficulties are encountered with peptides containing basic or aromatic amino-acids, both of which are present in fraction *B*.

hydrolysate was then transferred to a column of Amberlite IR-4B (British Drug Houses Ltd.) which had been equilibrated to pH 3.3, as described by Consden, Gordon & Martin (1948). The column, which was 10 cm. high and had a diameter of 1 cm., was washed with 200 ml. dil. HCl at pH 3.3, and the effluent, which contained neutral and basic peptides, was taken to dryness. The acidic peptides were eluted from the column with 100 ml. *N*-HCl and the solution taken to dryness and refractionated on a column of Amberlite IR-4B (10 cm. high) which had been equilibrated to pH 2.6. This column was developed with dil. HCl at pH 2.6. The first 100 ml. of effluent contained mainly peptides of aspartic and glutamic acids and was referred to as fraction *B1β*; the peptides of cysteic acid were collected in the next 200 ml. of effluent, which was fraction *B1α*. A further fraction could be eluted from the column with *N*-HCl. This contained chiefly free cysteic acid together with a few peptides.

The mixture of neutral and basic peptides obtained from the Amberlite-pH 3.3 column was dissolved in a small volume of 5% (w/v) aqueous acetic acid and transferred to a column of charcoal (activated charcoal B.D.H. previously washed with 20% acetic acid), of diameter 1 cm. and height

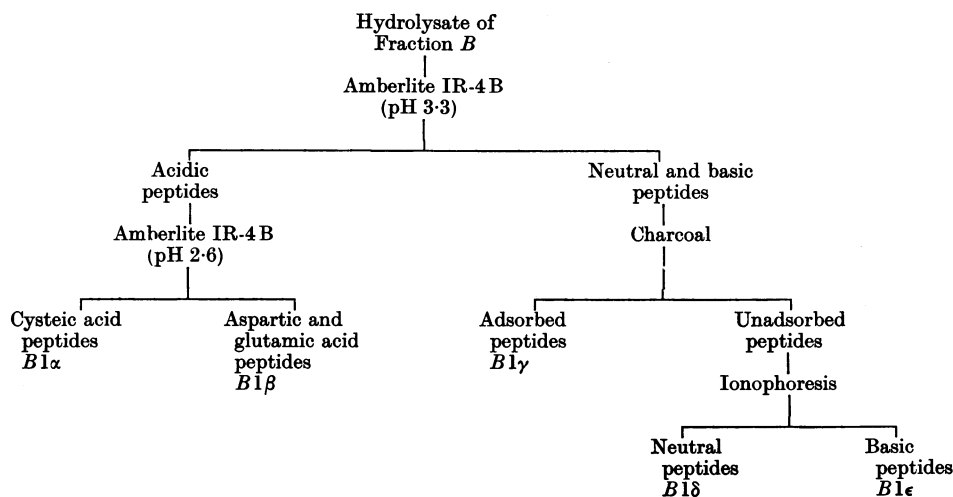


Fig. 1. Preliminary group separations in Exp. B1.

A number of experiments have been carried out using different methods both with the object of obtaining different peptides and also to investigate the various techniques. Three such experiments with hydrolysates made with 11*N*-HCl are described below. Probably none of them represents the most satisfactory method of fractionating the mixture, but they do indicate the types of separations that can be obtained. The first experiment may be regarded as a preliminary one and the others, in which fewer fractions were taken, were largely confirmatory. Since sufficient peptides were not obtained from the 11*N*-HCl hydrolysates to work out the entire sequence of amino-acids for fraction *B*, other methods of hydrolysis were studied. Experiments employing dilute acid (*B4*) and alkali (*B5*) for hydrolysis are also described.

Experiment B1. Fig. 1 summarizes the various stages of the fractionation used in this experiment. Fraction *B* (200 mg.) was hydrolysed for 4 days at 37° in 11*N*-HCl, and excess HCl was removed by evaporation *in vacuo*. The

10 cm., which was washed with 150 ml. 5% acetic acid. The effluent containing unadsorbed peptides was collected and taken to dryness. A further adsorbed fraction (*B1γ*) was obtained by washing the charcoal with a solution of 5% (w/v) phenol in 20% (w/v) aqueous acetic acid. The object of this treatment was to separate a fraction containing peptides of aromatic amino-acids (Tiselius, Drake & Hagdahl, 1947). In fact most of these peptides had already been lost by adsorption on the Amberlite, and this fraction contained other more strongly adsorbed peptides. Nevertheless, quite a valuable fractionation was obtained. The unadsorbed fraction was finally separated into a neutral (*B1δ*) and a basic (*B1ε*) fraction by ionophoresis in a simple three-compartment cell of the type used by Macpherson (1946).

Experiment B2. In this experiment we simply separated the hydrolysate into three fractions using a four-compartment ionophoresis cell that was designed by Dr R. L. M. Synge (1951) and which he kindly described to us. This is essentially the cell of Theorell & Åkeson (1942), to

which is added an extra anode compartment containing H_2SO_4 as in the desalting apparatus described by Consden, Gordon & Martin (1947). In this way the acid fraction is prevented from coming into contact with the anode, where decompositions are likely to take place, especially if chlorine is being evolved.

Fig. 2 shows a diagram of the cell as used in the present experiment. It differs only in certain details from that of Syngé. Membrane *a* is of cellophan, and membranes *b* and *c* of formolized gelatin (Syngé has membrane *b* of formolized sheep-skin parchment and *c* of cellophan). The cell was connected to 200 V. d.c. mains with a 40 W. filament lamp in series and was cooled during the run.

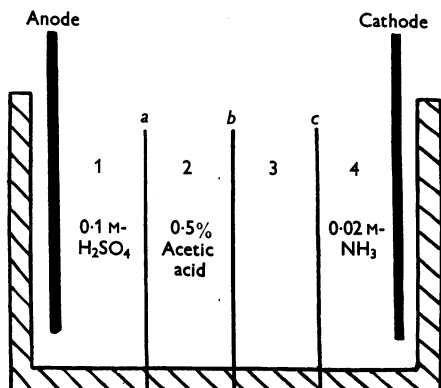


Fig. 2. Diagram of four-compartment ionophoresis cell. (For explanation see text.)

Fraction *B* (70 mg.) was hydrolysed 3 days at 37° in 11 *N*-HCl. After removal of excess HCl, the hydrolysate was adjusted to pH 5 and put in compartment 3 of the above cell. On applying the voltage an initial current of 110 ma. was produced and this rose to a maximum of 170 ma. and then slowly dropped. The pH of compartment 3 (which was stirred during the run) was maintained at about pH 4.5 (green to bromocresol green) by the addition of NH_3 . It is necessary to add considerable amounts of NH_3 which leads to an increase in the pH of compartment 4, which may prevent the weakly basic histidine peptides from entering this compartment. It is therefore necessary to change the aqueous NH_3 solution in compartment 4 twice during the run. This also prevents some decomposition of arginine peptides, which seems to occur on prolonged ionophoresis. The separation was continued for about 1 hr. when the current had dropped to 25–30 ma. The contents of compartments 2, 3 and 4 were collected and taken to dryness. The acidic fraction from compartment 2 is referred to as *B2α*, the neutral one from compartment 3 as *B2β* and the basic one from compartment 4 as *B2γ*.

Experiment B3. Fraction *B* (113 mg.) was hydrolysed in 11 *N*-HCl at 37° for 3 days. After removal of the HCl *in vacuo*, the residue was dissolved in 5 ml. water and shaken with 100 mg. charcoal for 20 min. The charcoal was filtered off and washed with 10 ml. 5% (w/v) aqueous acetic acid. The filtrate was then fractionated into acidic (*B3α*), neutral (*B3β*) and basic (*B3γ*) fractions in the four-compartment cell as described in Exp. B2. A further fraction (*B3δ*) was eluted from the charcoal with 20 ml. 5% phenol in 20% acetic acid.

Experiment B4. Fraction *B* (50 mg.) was boiled for 24 hr. under reflux with 0.1 *N*-HCl (50 ml.). After removal of the HCl *in vacuo* the residue was subjected to ionophoresis as described in Exp. B2, and acidic (*B4α*), neutral (*B4β*) and basic (*B4γ*) fractions were obtained.

Experiment B5. Fraction *B* (40 mg.) was heated at 105° in a sealed tube for 5 hr. with 5 ml. 0.2 *N*-NaOH. After neutralization with H_2SO_4 the mixture was desalted according to Consden, Gordon & Martin (1947). It was then subjected to ionophoresis as above.

Paper chromatography of peptides

Samples (corresponding to about 10 mg. of the original fraction *B*) of the various fractions obtained after group separations were subjected to two-dimensional chromatography on Whatman no. 4 filter paper. The various solvent systems used are described below:

Phenol-0.3% NH_3 . 0.3% (w/v) Aqueous NH_3 (1 l.) was put in the bottom of the chromatography tank. Phenol saturated with water was used in the trough, and the tank was filled with coal gas.

Cresol-0.03% NH_3 . As above, but containing 1 l. 0.03% (w/v) aqueous NH_3 and developed with *m*-cresol.

Collidine. The collidine used in this work gave considerably lower R_F values for the amino-acids than that used by Consden, Gordon & Martin (1944, see Dent, 1948).

Butanol-acetic acid. Prepared according to Partridge (1948) by mixing 5 vol. water, 4 vol. *n*-butanol and 1 vol. acetic acid. The mixture was usually kept a few days before use, and as no special precautions were taken to ensure that all samples were equally esterified the R_F values were not always constant and consequently have not been recorded. However, in Figs. 5–16, the distances of the peptide spots from the origin are given.

Other solvents used were prepared as described by Consden *et al.* (1944).

As a rule the chromatograms were developed first with phenol or cresol in the long direction of the paper and in the second direction with another solvent (usually butanol-acetic acid). If phenol was used as the second solvent, the fluorescence of the dried spots could not be seen. For certain slow-moving peptides it was necessary to run the solvent considerably farther than the end of the filter paper. This was accomplished by attaching a second folded sheet of filter paper to the front of the chromatogram. Where this was done the number of lengths of filter paper through which the solvent ran is indicated in Figs. 5–16. After development, the chromatograms were dried, heated at 100° for 15–30 min. to bring out the fluorescence of the spots (Phillips, 1948) and examined under the ultraviolet lamp. The fluorescent areas were marked on the paper. As observed by Jones (1949), the fluorescence was found to depend to a certain extent on the particular batch of filter paper used. In order to obtain the maximum effect it was necessary to avoid any contact of the papers with phenolic vapours as, for instance, in the drier. Where good fluorescent spots were seen they could be cut out directly without treatment with ninhydrin. Usually at least one replicate sheet was treated with ninhydrin to locate the peptides, and this could be used as a reference for cutting out the other sheets on which fluorescent spots were marked. 0.025% (w/v) Ninhydrin in water-saturated *n*-butanol was used to spray the sheets. This gave a sufficient colour development with very little destruction of the peptide, so that the eluates from the ninhydrin-treated spots could be used to

identify the amino-acids produced on hydrolysis. In certain cases, depending on the amino-acids concerned, there appeared to be a slight destruction of the *N*-terminal residue, but this was usually negligible. Where the peptide was to be deaminated or treated with fluorodinitrobenzene (FDNB), the use of ninhydrin was avoided and eluates from spots cut from untreated sheets were used. It was found that, if eluates from spots which had been treated with 0.025% ninhydrin were subjected to deamination, all the amino-acids present were almost completely destroyed.

Elution of peptides

For the elution of the peptides from the paper a combination of the methods of Consden, Gordon & Martin (1947) and of Dent (1947) was found convenient.

The spots were cut out and the cut trimmed to a point at one end. The other end was held between two microscope slides, which were placed in a trough of water as shown in Fig. 3. The solution passed into a capillary tube which was

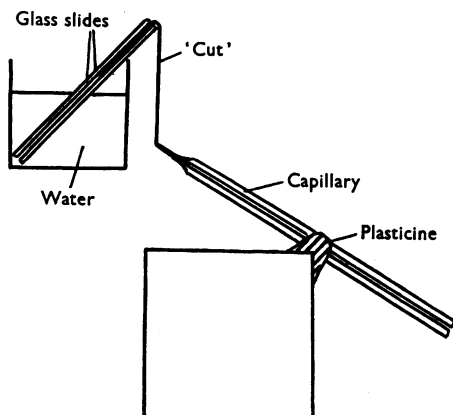


Fig. 3. Method of eluting peptides from filter paper.

held on a wooden block by a piece of plasticine. When sufficient liquid had collected, the contents were transferred to a polythene strip and dried *in vacuo*.

Identification of amino-acid residues present in peptides

Hydrolysis of the peptides was carried out with 5.7 *N*-HCl in capillary tubes as described by Consden, Gordon & Martin (1947), and the resulting amino-acids identified by paper chromatography. Fraction *B* contains the amino-acids listed in Table 14. It was shown by chromatography with benzyl alcohol that isoleucine is absent.

On account of the limited chromatographic apparatus available, and the labour involved, it was desirable to use one-dimensional chromatograms as far as possible to identify the amino-acids present. In the earlier experiments phenol-0.3% NH_3 was used for the preliminary identification of all peptide hydrolysates. Cysteic acid, aspartic acid, glutamic acid, serine, glycine, threonine and alanine can all be unequivocally identified in this solvent. The other faster moving amino-acids cannot be so easily distinguished especially if several of them are present. They may be more

satisfactorily identified if phenol-0.03% NH_3 is used. The basic amino-acids move slower on this solvent and may frequently be recognized. The identification of all these faster moving amino-acids was always confirmed using some other solvent system or on a two-dimensional chromatogram.

For the complete analysis of a peptide two-dimensional chromatograms were run on half sheets (46 × 28 cm.) of Whatman no. 1 filter paper, which were developed in the long direction with phenol-0.3% NH_3 and in the short with butanol-acetic acid. All the amino-acids present in fraction *B* may be recognized on such a chromatogram (Fig. 4*b*).

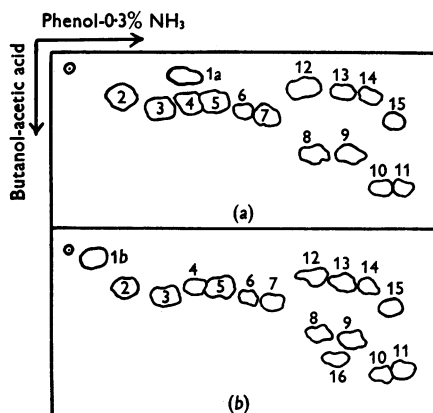


Fig. 4. Chromatograms of hydrolysates of insulin (*a*) and fraction *B* (*b*) run on half sheets of Whatman no. 1 filter paper. The identities of the spots are as follows: 1*a*, cystine; 1*b*, cysteic acid; 2, aspartic acid; 3, glutamic acid; 4, serine; 5, glycine; 6, threonine; 7, alanine; 8, tyrosine; 9, valine; 10, leucine; 11, phenylalanine; 12, histidine; 13, lysine; 14, arginine; 15, proline; 16, TyrX.

On hydrolysis of certain aromatic peptides a new spot was observed moving slightly faster than tyrosine in phenol and almost as fast as leucine in butanol-acetic acid. It gave a positive reaction with diazotized sulphanic acid. In many cases pairs of peptides were found, one containing tyrosine and the other this new amino-acid. It is thus considered to be a breakdown product of tyrosine produced during the oxidation with performic acid, and will be referred to as TyrX. Fig. 4*b* shows a chromatogram of a hydrolysate of fraction *B* and Fig. 4*a* one of a hydrolysate of insulin. The TyrX spot (16) can be seen only in Fig. 4*b* below the tyrosine and valine spots, thus showing that it has been formed during the preparation of fraction *B*.

Since histidine gives a relatively weak spot with ninhydrin, it is possible to overlook it in certain peptides. The following modification of the Pauli test which avoids the use of aqueous solutions was found to be suitable for the identification of histidine and histidine peptides on chromatograms.

A mixture of equal volumes of 1% (w/v) *p*-anisidine in ethanolic 0.11 *N*-HCl and 10% (v/v) amyl nitrite in ethanol was allowed to stand a few minutes and then sprayed on to the paper. The sheets were dried in air and the colour developed with NH_3 vapour or by spraying on ethanolic KOH solution. Histidine and histidine peptides give a dark-red spot, whereas tyrosine gives a faint orange colour.

Determination of the structure of peptides

Deamination. In order to determine the *N*-terminal residues of the peptides we have deaminated them with nitrosyl chloride as described by Consden, Gordon & Martin (1947). In some experiments the reaction was carried out at 37° for 10 min., in others at room temperature for 30 min. No difference was noticed between the two procedures. For most peptides the method gave satisfactory results, though it was rarely possible to destroy the *N*-terminal residue completely. Certain amino-acids such as leucine and glutamic acid seemed to be especially resistant to this treatment, and if these were present as *N*-terminal residues, clear-cut deamination could rarely be obtained. Results with peptides containing tyrosine or arginine were also difficult to interpret.

DNP method. For most of the peptides the *N*-terminal residue has also been determined by the DNP technique, which usually gives clear-cut results, though it is rather more laborious. In some cases the DNP-derivative of the terminal residue was identified directly by chromatography using the micro method previously described (Sanger, 1949*b*). As this requires rather more material than is usually available from a paper chromatogram, it is often convenient to use an alternative procedure, and to identify the amino-acids present after hydrolysis of the DNP-peptide. Although traces of free amino-acids are formed on hydrolysis of DNP amino-acids (Consden, Gordon, Martin & Synge, 1947), it was found that they are sufficiently stable to ensure that the terminal residue is almost completely absent. The results are usually much more clear-cut than when the deamination method is used.

In this method the peptide is treated with FDNB in the usual way and the excess reagent removed by extraction of the alkaline solution with ether. It is then necessary to separate the DNP-peptide from the salt present. This may be done by extracting it from acid solution into ethyl acetate, or if it cannot be extracted by adsorbing it on a talc column from acid or neutral solution (Sanger, 1949*b*). The DNP peptide is then hydrolysed for 20 hr. with 5·7*N*-HCl at 105° and the amino-acids remaining identified. In the case of peptides containing glycine as *N*-terminal residue a certain amount of breakdown of the DNP-glycine to free glycine was observed so that the results were not so clear. Better results were obtained if a shorter time of hydrolysis was used. Lysine, tyrosine and histidine, if present in a peptide with their amino groups blocked, are converted to their monoDNP derivatives which can be identified on paper chromatograms. Table 2 shows *R_F* values for these with butanol-acetic acid which had been prepared 24 hr. before use. The values for α -*N*-DNP-arginine, DNP-cysteic acid and for valine and leucine are given for comparison. Although *O*-DNP-tyrosine and *im*-DNP-histidine are colourless, they show up as dark

spots if the papers are examined in ultraviolet light. This effect would seem to be due to quenching the fluorescence of the paper. Peptides containing *im*-DNP-histidine may be successfully fractionated on paper chromatograms with solvents such as collidine and pentanol.

RESULTS

The results obtained with the various fractions are given in Figs. 5–16 and Tables 3–13.

Figs. 5–16 are diagrams of the two-dimensional chromatograms, from which the peptide spots were eluted, and Tables 3–13 give the results obtained with the eluted peptides. In column 2 of these tables are listed the amino-acids identified on complete hydrolysis, and the approximate strength is indicated in column 3. The \times 's have the same significance as in the paper of Consden *et al.* (1949), and ? represents a very weak spot. Column 4 shows the strength of the amino-acids remaining after deamination with NOCl and hydrolysis of the peptide and column 5 that remaining after hydrolysis of the DNP peptide. In the last column are given the structures of the peptides, as far as they can be deduced from the data given in the tables. As far as possible all the results with one peptide are given together, although they were often obtained in different experiments, and sometimes from the same peptide in a different fraction. Most of the results recorded have been obtained in several experiments. Where deamination or the DNP method has been used only the most satisfactory and clear-cut results have been recorded. While such a procedure does not necessarily give a true indication of the efficiencies of the methods it is probably justified, as much more can be learnt from such a result than from an indefinite one. This is especially true for the deamination technique where several experiments were usually carried out on each peptide.

Besides the peptides shown in Fig. 5 and Table 3, fraction *B*1 α contained traces of $\text{CySO}_3\text{H} \cdot \text{Ala}$ and $\text{Ser} \cdot \text{Val} \cdot \text{CySO}_3\text{H}$. These peptides were probably derived from the small amount of fraction *A* contained in the preparation of fraction *B*, since they were found in very high concentration in fraction *A*. The presence of $\text{Ser} \cdot \text{Val} \cdot \text{CySO}_3\text{H}$ is suggested by a trace of serine that was sometimes noticed in spot 3. With this fraction there was frequently overlapping between spots 3 and 4 and between spots 5 and 6. When the eluate from the Amberlite-pH 2·6 column was collected in several successive fractions, a slower moving fraction was obtained which contained peptides 4 and 6 without the corresponding peptides containing glycine. The smaller peptides are retained more strongly on the resin than the larger.

Two chromatograms of fraction *B*1 γ are illustrated in Figs. 7 and 8, since certain peptides are separated better using cresol and others with phenol. The combined results are given in Table 5.

Table 2. *R_F* values of water-soluble DNP derivatives

(Whatman no. 4 filter paper; butanol-acetic acid.)

	<i>R_F</i>	Colour of spot	Ninhydrin reaction
<i>O</i> -DNP-Tyrosine	0·84	Colourless	Purple
α - <i>N</i> -DNP-Arginine	0·81	Yellow	No reaction
ϵ - <i>N</i> -DNP-Lysine	0·77	Yellow	Brown
Leucine	0·67	—	—
<i>im</i> -DNP-Histidine	0·57	Faint yellow	Brown
Valine	0·49	—	—
DNP-Cysteic acid	0·42	Yellow	—

The results with fraction *B1ε* are recorded with the results on fraction *B2γ*, which gave a more complete mixture of basic peptides. *B1ε* was identical with *B2γ* except that

The results with fractions *B2α* and *B2β* (Figs. 10–12; Tables 7 and 8) which are largely confirmatory, are included to show the type of separation that can be obtained with mixtures of about twenty peptides, and also to show that very few peptides were overlooked in the preliminary experiments. The cresol chromatogram of fraction *B2α* (Fig. 11) is included to demonstrate the positions of the fast-moving spots 21 and 22, which did not appear on the phenol chromatogram.

A large number of solvents was studied for fractionating the basic peptides, most of which move fast on systems such as phenol-0.3% NH_3 and slow on butanol-acetic acid or collidine. Cresol-0.03% NH_3 was eventually found to give good separations (Fig. 13) and is to be recommended for the fractionation of basic peptides and also of larger peptides.

In one experiment fraction *B2γ* was subjected to a further ionophoresis to obtain a fraction containing only arginine peptides. This was carried out in the four-compartment cell (Fig. 2). Compartment 2 contained 1% (w/v) acetic acid and compartment 4 0.1M-aqueous NH_3 . Compartment 3 was maintained at a pink reaction to phenolphthalein. The maximum current was 120 ma. and dropped to 30 ma. The basic fraction (from compartment 4) contained only arginine and Arg.Gly (*B2γ5*). The fraction from compartment 3 contained traces of these besides the other basic peptides

Considerable difficulty was experienced in fractionating the fast-moving neutral peptides, which appeared in spots *B2β24* and *B2β25* in Exp. *B2*. The most satisfactory solvents

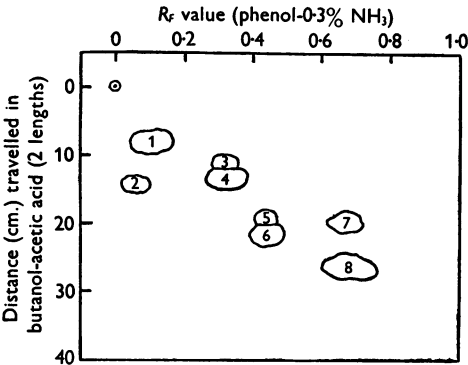


Fig. 5. Chromatogram of fraction *B1α* (see Table 3). In Figs. 5–16 the fraction was applied in the position marked by the small dotted circle, and the chromatogram run first in the direction represented horizontally, which, except in Fig. 15, was the longer dimension of the filter paper.

Ser.[His, Leu] (*B2γ7*) and His.Leu (*B2γ8*) were absent, since they had been adsorbed on charcoal and appeared in fraction *B1γ*.

Table 3. Peptides from fraction *B1α*

Spot no. (Fig. 5)	Amino-acids present	Strength of amino-acid after			Structure
		Hydrolysis	Deamination and hydrolysis	DNP treatment and hydrolysis	
1	CySO ₃ H Gly	x x x x x x x x	? x x x x	— x x x	CySO ₃ H. Gly
2	Asp Glu	x x			[Asp, Glu]
3	CySO ₃ H Gly Val	x x x x x		x x —	Val.[CySO ₃ H, Gly]
4	CySO ₃ H Val	x x x x x x	x x ?	x x —	Val.CySO ₃ H
5	CySO ₃ H Gly Leu	x x x	x ? —	x ? —	Leu.[CySO ₃ H, Gly]
6	CySO ₃ H Leu	x x x x x x	x x x x x	x x x —	Leu.CySO ₃ H
7	CySO ₃ H Gly Val Leu	x x x x x x x x	x x x ?	x x x x x x —	Leu.[CySO ₃ H, Gly, Val]
8	CySO ₃ H Val Leu	x x x x x x x x x x x x	x x x x x x x	x x x x x x —	Leu.[CySO ₃ H, Val]

Spot 1. Yellow colour with ninhydrin which on further heating became grey and then violet.
Spot 8. Eluates of spot 8 from two chromatograms were hydrolysed 4 days at 37° in 11N-HCl and investigated on a phenol-0.3% NH_3 chromatogram. Besides spots corresponding to cysteic acid, valine, leucine and the unchanged peptide there was a spot having R_f 0.32 which contained valine and cysteic acid and a spot moving faster than leucine which contained valine. Since valine is present in both these dipeptides it must be in the middle of peptide 8, whose structure is therefore Leu.Val.CySO₃H.

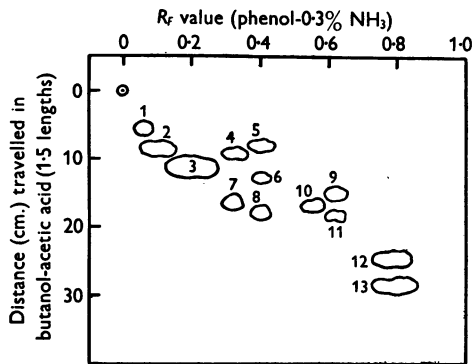


Fig. 6. Chromatogram of fraction *B1β* (see Table 4).

Table 4. *Peptides from fraction B1β*

Spot no. (Fig. 6)	Amino-acids present	Strength of amino-acid after			Structure
		Hydrolysis	Deamination and hydrolysis	DNP treatment and hydrolysis	
1	CySO ₃ H Gly	x x x x			CySO ₃ H.Gly (<i>B1α1</i>)
2	Asp	x x x x			
3	Glu	x x x x x			Glutamic acid
4	CySO ₃ H Gly Val	x x x x			Val.[CySO ₃ H, Gly] (<i>B1α3</i>)
5	CySO ₃ H Glu His Leu	x x x x			
6	CySO ₃ H Gly Leu	x x x			Leu.[CySO ₃ H, Gly] (<i>B1α5</i>)
7	Glu Ala	x x			
8	Asp Val	x x	? —	x —	Val.Asp
9	CySO ₃ H Gly Val Leu	x x x x			
10	Asp Glu Val	x x x x x x	x x —	x x x x —	Val.[Asp, Glu]
11	CySO ₃ H Val Leu	x x x			
12	Asp Glu Val Phe	x x x x x x x x x x x x x x x x		x x x x x x x x x x x x	Phe.[Asp, Glu, Val]
13	Asp Val Phe	x x x x x x x x x x x x		x x x x —	

Table 5. *Peptides from fraction B1 γ*

Spot no. (Figs. 7 and 8)	Amino-acids present	Strength of amino-acid after			Structure
		Hydrolysis	Deamination and hydrolysis	DNP treatment and hydrolysis	
1	Glu	x x		?	Glu.His
	His	x x		x*	
2	Glu	x x			[Glu, Gly, Arg]
	Gly	x x			
	Arg	x x			
3	Glu	x x x	x		Glu.Arg + spot 2
	Gly	x	?		
	Arg	x x x	x x x†		
4	CySO ₃ H	x x x	x x	x x	His.[CySO ₃ H, Leu]
	His	x x x	?	—	
	Leu	x x x	x x x	x x	
5	Thr	x x x x			Thr.[Ala, Lys, Pro] (B2γ6)
	Ala	x x x x			
	Lys	x x x x			
	Pro	x x x			
6	Asp	x		x	Phe.[Asp, Glu, His, Val]
	Glu	x		x	
	His	x		†	
	Val	x		x	
	Phe	x		—	
7	Glu	x x			[Glu, His, Leu]
	His	x x			
	Leu	x x			
8	Leu	x x			His.Leu (B2γ8)
	His	x x			
9	Glu	x			Val.Glu (B187)
	Val	x			
10	Glu	x	x		Val.[Ala, Glu]
	Ala	x	x		
	Val	x	—		
11	Tyr	x			Tyrosine
12a	Gly	x			Gly.Phe
	Phe	x			
12b	Asp	x x		x	See below
	Glu	x x		x	
	Gly	x x		?	
	Val	x x		x	
	Phe	x x x		x x	
13	Leu	x			Leucine + phenylalanine
	Phe	x x			
14	Glu	x x x			Leu.[Glu, Val] (B2β14)
	Val	x x x			
	Leu	x x x			
15	Glu	x			Leu.[Glu, Ala, Val] (B2β15)
	Ala	x			
	Val	x			
	Leu	x			
16	Ala	x x			Ala.Leu (B1811)
	Leu	x x			
17	Unidentified mixture of fast-moving peptides, containing Phe, Leu, Val, Tyr, Ala				

* DNP derivative.

† Arginine decomposition product (R_F , 0.73 with phenol).‡ *m*-DNP-Histidine not distinguished on this phenol chromatogram.

Spots 2 and 3 appear to be mixtures of a tripeptide [Glu, Gly, Arg] and a dipeptide [Glu, Arg], the latter moving slightly faster on butanol-acetic acid.

Spot 7, obtained from phenol chromatograms, sometimes contains serine probably due to the presence of Ser.[His, Leu] (B2 γ 7).

Spot 8 sometimes contains glutamic acid from spot 7.

Spot 12a gives a yellow colour with ninhydrin, suggesting glycine is terminal. The DNP treatment was done only on the mixture in spot 12b, which clearly demonstrates that glycine must be terminal.

Spot 12b is a mixture of 12a with another peptide, probably Phe.[Asp, Glu, Val] (B1 β 12). This latter did not appear on the cresol chromatogram (Fig. 8).

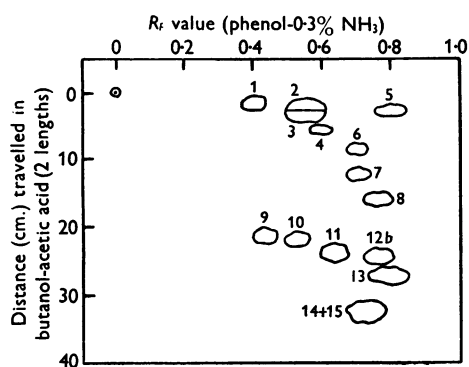


Fig. 7. Chromatogram of fraction *B1γ* using phenol-0.3% NH_3 (see Table 5).

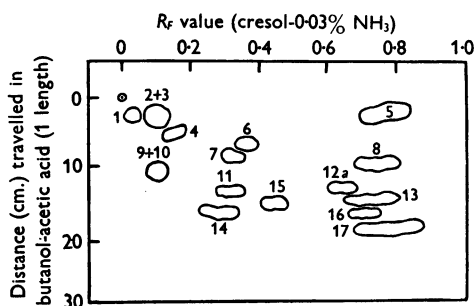


Fig. 8. Chromatogram of fraction *B1γ* using cresol-0.03% NH_3 (see Table 5).

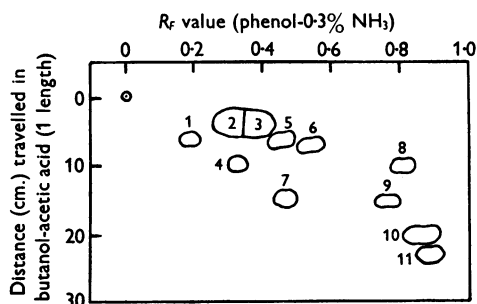


Fig. 9. Chromatogram of fraction *B1δ* (see Table 6).

Table 6. *Peptides from fraction B1δ*

Spot no. (Fig. 9)	Amino-acids present	Strength of amino-acid after			Structure
		Hydrolysis	Deamination and hydrolysis	DNP treatment and hydrolysis	
1	Glu	x	x	x	Gly. Glu
	Gly	x	—	?	
2	Ser	x x x			Serine
	Gly	x			
3	Gly	x x x x			Glycine
4	Glu	x	?		Glu. Ala
	Ala	x	x		
5	Thr	x			Threonine
6	Ala	x x x x			Alanine
7	Glu	x x	x x	x x	Val. Glu
	Val	x x	—	?	
8	Thr	x x	?	—	Thr. Pro
	Pro	x x	x x	x x x	
9	Val	x x			Valine
10	Leu	x x x x			Leucine
11	Leu	x x x x	x x x x	x x x x	Ala. Leu
	Ala	x x x x	x x	x	

Spot 1 gives a yellow colour with ninhydrin.

Spot 7 gives a very weak reaction with ninhydrin and a relatively strong fluorescence.

Spot 8 gives a strong yellow colour with ninhydrin.

Spot 10 is often contaminated with traces of serine, which is located in the slower moving parts of the spot. If *m*-cresol is used instead of phenol the leucine spot is pure and the serine can be found together with leucine in a slower moving faint spot (R_f , 0.57). This is probably Ser. Leu derived from fraction *A*, from which it can be obtained in large amounts.

Spot 11 is often contaminated with threonine, which we have not yet been able to account for satisfactorily.

were collidine *tert*-amyl alcohol (2-methylbutan-2-ol) and benzyl alcohol, but there seemed to be too many of these closely related peptides to give clear results. The best results were obtained with fraction *B3β* (Fig. 14, Table 10), from which some of the more strongly adsorbed peptides had been removed by charcoal treatment. The very large number of these peptides containing tyrosine, valine, leucine and phenylalanine is partly due to the partial breakdown of tyrosine to TyrX, which means that two spots are obtained for each tyrosine peptide. It also suggests that these non-polar residues are grouped together in the fraction *B* chains.

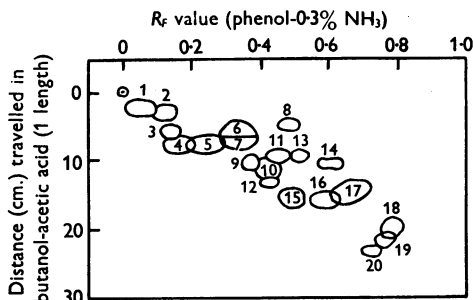


Fig. 10. Chromatogram of fraction *B2α* using phenol-0.3% NH_3 (see Table 7).

The results with the other fractions of Exp. *B3* are not recorded, since they were essentially the same as the *B2* fractions. Fraction *B3δ*, which was adsorbed on charcoal, contained mainly [CySO₃H, Tyr, Val, Leu] (*B2α21*) and Phe.[Asp, Glu, Val] (*B1β12*) and fast-moving peptides containing tyrosine, valine, leucine and phenylalanine.

The peptides found in fraction *B4α* are listed in Table 11. Fractions *B4β* and *B4γ* yielded a few new peptides, which are recorded in Table 12. Other peptides are not recorded. Spot *B4β1* was obtained from a cresol-0.03% NH_3 /butanol-acetic acid chromatogram and ran at the same rate in both solvents as His.[CySO₃H, Leu] (Peptide *B2β6*, Fig. 12). Peptide *B4β2* was a fast-moving spot on a collidine, *tert*-amyl alcohol chromatogram. The peptides from *B4γ* were obtained from the chromatogram illustrated in Fig. 15.

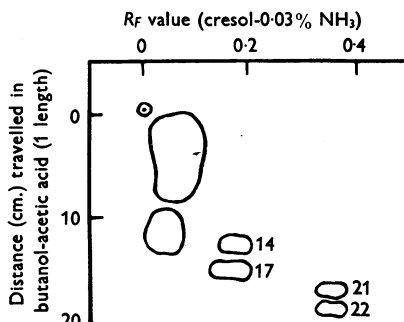


Fig. 11. Chromatogram of fraction *B2α* using cresol-0.03% NH_3 (see Table 7).

Fraction *B5α* contained very little material and was not investigated in detail. *B5β* contained free glycine and alanine (but no other free amino-acids), Gly.[Glu, Orn] (*B5γ1*), Gly. Glu (*B1δ1*) and a large amount of fast-moving peptides. There was also a yellow spot ($R_F=0.55$ in cresol-0.03% NH_3 and a little faster than alanine on butanol-acetic acid) which gave glycine and proline on hydrolysis. It

Table 7. Peptides from fraction *B2α*

Spot no. (Figs. 10 and 11)	Strength	Structure	Identical with
1	x x x	Cysteic acid	—
2	x x x	CySO ₃ H. Gly	<i>B1α1</i>
3	x x	Aspartic acid	—
4	x x	Gly. Glu	<i>B1δ1</i>
5	x x x	Glutamic acid	—
6	x x	Val.[CySO ₃ H, Gly]	<i>B1α3</i>
7	x	Val. CySO ₃ H	<i>B1α4</i>
8	x	[CySO ₃ H, Glu, His, Leu]	<i>B1β5</i>
9	x x	Glu. Ala	<i>B1δ4</i>
10	x x x	Leu. CySO ₃ H	<i>B1α6</i>
11	x x x	Leu.[CySO ₃ H, Gly]	<i>B1α5</i>
12	x x	Val. Asp	<i>B1β8</i>
13	x	Alanine	—
14	x x x	Leu.[CySO ₃ H, Gly, Val]	<i>B1α7</i>
15	x x x x	Val. Glu	<i>B1δ7</i>
16	x	Val.[Ala, Glu]	<i>B1γ10</i>
17	x x x x	Leu.[CySO ₃ H, Val]	<i>B1α8</i>
18	x x	Phe.[Asp, Glu, Val]	<i>B1β12</i>
19	x	Leu.[Glu, Ala, Val]	<i>B2β15</i>
20	x x	Leu.[Glu, Val]	<i>B2β14</i>
21	x	[CySO ₃ H, TyrX, Val, Leu]	—
22	x	[CySO ₃ H, TyrXX, Val, Leu]	—

Spot 8 sometimes contains a trace of glycine.

Spot 22. Besides cysteic acid, valine and leucine, a new spot was present that had R_F , 0.53 (phenol-0.3% NH_3) and moved as fast as leucine on butanol-acetic acid. It was not detected in any other peptides and is assumed to be another breakdown product of tyrosine (TyrXX).

is assumed that this was formed by the action of the alkali on Thr.Pro (B188). Wieland & Wirth (1949) have shown that glycine is formed from threonine by the action of alkali.

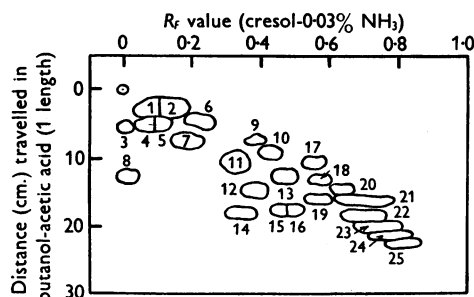


Fig. 12. Chromatogram of fraction B2β (see Table 8).

The results with fraction B5γ are recorded in Fig. 16 and Table 13. Besides the peptides shown there were a few unidentified spots that moved faster on butanol-acetic acid. The amino-acid referred to as Orn moved at the same R_f values as ornithine in phenol and butanol-acetic acid and is assumed to be a breakdown product of arginine.

DISCUSSION

Methods of hydrolysis

The present results support the conclusion that concentrated hydrochloric acid at low temperatures is the most generally useful method of partial

hydrolysis (Synge, 1943). No apparent destruction of amino-acids occurs, and a wide variety of lower peptides is produced. The disadvantage of using only one method is that it is difficult to obtain peptides which contain the most labile bonds, and which are required to work out a complete sequence of amino-acids. The bonds that are most labile under these conditions are those involving the amino groups of serine and threonine residues (Desnuelle & Casal, 1948) and no peptides have been found containing these bonds intact. Bonds involving glycine residues are also somewhat labile, but some such peptides were obtained. Other methods of hydrolysis were therefore studied in the hope of supplementing the results from the hydrolysates made with concentrated hydrochloric acid and also in the hope of obtaining a rather less complex mixture of the fast-moving neutral peptides, suitable for fractionation. While neither of these objectives was attained, a few new peptides were detected.

In studying the terminal peptides (Sanger, 1949*b*) it was found that glycine peptides were relatively more stable to dilute acid than to concentrated acid. This is also apparent from the present work. If fraction B4α (Table 11) is compared with fraction B1α (Table 3) it will be seen that the dilute-acid hydrolysate contains a higher proportion of the peptides in which the CySO_3H .Gly bond is intact. Similarly, $[\text{CySO}_3\text{H}, \text{Gly}, \text{His}, \text{Leu}]$ (B4β1) occurs in

Table 8. Peptides from fraction B2β

Spot no. (Fig. 12)	Strength	Structure	Identical with
1	x x	Glu. His	B1γ1
2	x x x x	Glu. Arg + [Glu, Gly, Arg]	B1γ2 + B1γ3
3	x	Gly. Glu	B1δ1
4	x x x x	Serine + glycine	—
5	x x x	Glycine + threonine	—
6	x x	His. [CySO ₃ H, Leu]	B1γ4
7	x x x x	Alanine	—
8	x	Val. Glu + Val. [Ala, Glu]	B1γ9 + B1γ10
9	x	Phe [Asp, Glu, His, Val]	B1γ6
10	x x	[Glu, Ser, His, Val, Leu]	B4γ3
11	x x	Tyrosine	—
12	x	TyrX	—
13	x x	Valine	—
14	x x x x	Leu. [Glu, Val]	—
15	x x	Leu. [Glu, Ala, Val]	—
16	x x	Phe. [Asp, Val]	B1β13
17	x x x	Thr. Pro	B188
18	x x	[Glu, His(?)Val, Leu]	—
19	x x	Phe. [Asp, Glu, Val]	B1β12
20	x x	Gly. Phe	B1γ12a
21	x x x x	Phenylalanine + leucine	—
22	x x x x	Ala. Leu	B1811
23	x x x	Ala. [Tyr, Leu]	B3β9

Spot 6 contains a trace of glycine, probably due to $[\text{His}, \text{Leu}, \text{CySO}_3\text{H}, \text{Gly}]$ (B4β1).

Spot 10. Glutamic acid, histidine and leucine are slightly stronger so probably contains peptide B1γ7 ([Glu, His, Leu]).

Spot 14. After DNP treatment only valine and glutamic acid were present.

Spot 15. N-terminal residue determined on mixture in B3β4.

Spot 18. The histidine was not clearly distinguished on this phenol chromatogram as it comes very close to the valine spot, but it is probably significant since the peptide cannot be $[\text{Glu}, \text{Val}, \text{Leu}]$ which is in spot 14.

Spots 24, 25. Unidentified mixtures of peptides containing tyrosine, valine, leucine and phenylalanine.

the dilute-acid hydrolysate, whereas the corresponding peptide without glycine (*B*2 β 6) is obtained with concentrated hydrochloric acid hydrolysis. The presence of several new histidine peptides in fraction

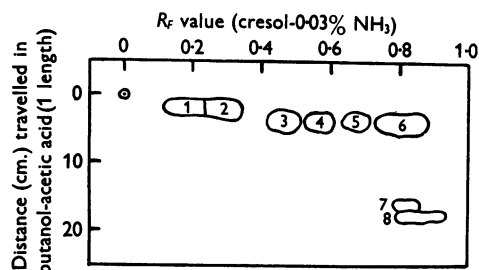


Fig. 13. Chromatogram of fraction *B*2 γ (see Table 9).

*B*4 γ (Table 12) suggests that bonds involving histidine may also be relatively more stable to dilute acid. The bonds involving the amino groups of serine and threonine, however, appear to be labile to dilute acid also, and no peptides were obtained in

which they were intact. The large amount of free aspartic acid (*B*4 α 2), compared with only traces of free cysteic acid (possibly present in spot *B*4 α 1), is interesting in view of the results of Partridge & Davis (1950), who demonstrated a preferential release of aspartic acid when proteins were boiled with weak acids.

The hydrolysis in alkali led to considerable destruction of the amino-acids, apparently while they were still in peptide linkage, which greatly complicates the results. Wieland & Wirth (1949) found no destruction of serine or threonine on heating 12 hr. with saturated barium hydroxide, conditions which are more violent than those used here. However, the residues seem considerably more labile when in peptide linkage. No free serine or threonine was detected, and the Gly-Pro was almost certainly derived from Thr.Pro. Arginine appears to be converted almost quantitatively to ornithine since no free arginine was detected, whereas free ornithine and a peptide containing ornithine (*B*5 γ 1) were present.

Table 9. *Peptides from fraction B2γ*

Spot no. (Fig. 13)	Amino-acids present	Strength of amino-acid after			Structure
		Hydrolysis	Deamination and hydrolysis	DNP treatment and hydrolysis	
1	Ser	x x x			2 + lysine
	Lys	x			
	His	x x x			
2	Ser	x x x x	x	—	Ser.His
	His	x x x x	x x x	x x *	
3	His	x x			Arginine + histidine
	Arg	x x x			
4	Ala	x x x		x x	Lys.Ala
	Lys	x x x		—	
5	Gly	x x x		x x	Arg.Gly
	Arg	x x x		—	
6	Thr	x x x x x		x	Thr.[Ala, Lys, Pro]
	Ala	x x x x x		x x x	
	Lys	x x x x x		x x x *	
	Pro	x x x x x		x x x	
7	Ser	x x		—	Ser.[His, Leu]
	His	x x		x *	
	Leu	x x		x	
8	His	x x x	?	—	His.Leu
	Leu	x x x	x x	x x	

* DNP derivative.

This fraction also contained traces of free neutral amino-acids, which we have always found in the basic fractions from the ionophoresis. These are not recorded.

Spot 1. Part of the Ser.His spot contains free lysine. These two are well separated if 0.3% NH_3 is used with the cresol, when the lysine travels much faster. However, the other peptides are badly separated and only Ser.His appears as a single pure spot.

Spot 4. Hydrolysis of the DNP derivative gave bis-DNP-lysine.

Spot 5. The DNP derivative of this basic peptide could be extracted (though not rapidly) into ethyl acetate from acid solution, in contrast to DNP-arginine. It was not well adsorbed on talc from acid solution, but could be adsorbed better from neutral solution.

Spot 6. Contains traces of glycine which has probably arisen by partial breakdown of the *N*-terminal threonine.

Methods of group separation

Probably none of the experiments described represent the most satisfactory procedure for fractionating the mixture concerned. Exp. B1, in which most of the aromatic peptides were lost by adsorption on the resin, was designed primarily to separate cysteic acid peptides and it was only later decided to study the other fractions.

The four-compartment cell of Syngé was found to give very satisfactory separations into acidic, basic and neutral peptides with only a little overlapping. Under the conditions described many of the simple peptides of glutamic acid were found in the neutral fraction (B2β) as well as in the acidic fraction (B2α). By maintaining the specimen compartment at a slightly higher pH this overlapping could be avoided and all the glutamic acid peptides were then found in the acidic fraction, but under these conditions the neutral fraction contained histidine peptides.

The ion-exchange resin gave a good separation of peptides of cysteic acid, which are less strongly adsorbed than free cysteic acid. No clear separation

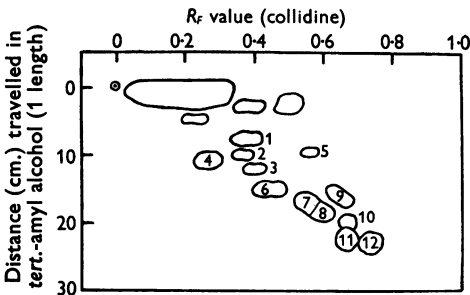


Fig. 14. Chromatogram of fraction B3β (see Table 10).

of peptides containing aromatic residues was obtained by adsorption on charcoal. Several other peptides were also strongly adsorbed. Most of these

Table 10. Peptides from fraction B3β

Spot no. (Fig. 14)	Amino-acids present	Strength of amino-acid after		Structure
		Hydrolysis	DNP treatment and hydrolysis	
1	Asp	x x	x	Phe.[Asp, Glu, Val] (B1β12) + Gly.Phe (B1γ12α)
	Glu	x x	x	
	Gly	x x	?	
	Val	x x	x	
	Phe	x x x	x	
2	Asp	x x		Phe.[Asp, Val] (B1β13)
	Val	x x		
	Phe	x x		
3	Ala	x x x x x		Ala.Leu (B1δ11)
	Leu	x x x x x		
4	Glu	x x x x	x x x	Leu.[Glu, Val] (B2β14) + Leu.[Glu, Ala, Val] (B2β15)
	Ala	x x	x	
	Val	x x x x	x x x	
	Leu	x x x x	?	
5	TyrX	x x		TyrX
6	Phe	x x x		Phenylalanine + leucine
	Leu	x x x x		
7	Val	x x x	x x x	Leu.Val
	Leu	x x x	x	
8	Val	x x		[Val, Phe] (B4β2) + 7
	Leu	x x		
	Phe	x		
9	Ala	x x	—	Ala.[Tyr, Leu]
	Tyr	x x	?*	
	Leu	x x x	x	
10	TyrX + Tyr	x x	?	[TyrX, Leu] [Tyr, Leu]
	Leu	x x	x	
11	Tyr	?		Mixture
	Val	x		
	Leu	x x		
	Phe	x x		
12	Tyr	x		[Tyr, Val, Leu]
	Val	x x		
	Leu	x x		

* DNP derivative.

Table 11. *Peptides from fraction B4α*

Peptide	Strength	Structure	Identical with
1	x x x	CySO ₃ H. Gly	B1α1
2	x x x x	Aspartic acid	—
3	x x x	Glutamic acid	—
4	x x	Val.[CySO ₃ H, Gly]	B1α3
5	x x	Leu.[CySO ₃ H, Gly]	B1α5
6	x	Leu.CySO ₃ H	B1α6
7	x x	Leu.[CySO ₃ H, Gly, Val]	B1α7
8	x	Leu.[Val, CySO ₃ H]	B1α8

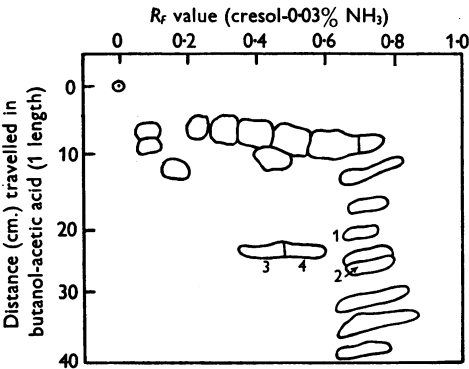


Fig. 15. Chromatogram of fraction *B4γ*. The chromatogram was run first down the short direction of the filter paper with cresol-0.03 % NH₃ and then down the long direction with butanol-acetic acid (see Table 12).

Table 12. *Peptides from fractions B4β and B4γ*

Peptide	Amino-acids present	Strength of amino-acid after		Structure
		Hydrolysis	DNP treatment and hydrolysis	
β1	CySO ₃ H	x x x	}	[CySO ₃ H, Gly, His, Leu]
	Gly	x x x		
	His	x x x		
	Leu	x x x		
β2	Val	x x x	}	[Val, Phe]
	Phe	x x		
γ1*	Ser	x x	}	Ser.[His, Leu] (<i>B2γ7</i>)
	His	x x		
	Leu	x x x		
γ2*	Ser	x	}	Ser.[His, Val, Leu]
	His	x x		
	Val	x x		
	Leu	x x		
γ3*	Glu	x x	}	[Glu, Ser, His, Val, Leu]
	Ser	x x		
	His	x x		
	Val	x x x		
	Leu	x x x		
γ4*	Glu	x x	}	Ser.[Glu, Ala, His, Val, Leu]
	Ser	x		
	Ala	x x		
	His	x x		
	Val	x x		
	Leu	x x		

* Spot nos. as in Fig. 15. These gave yellow colour with ninhydrin.
† Run on a phenol chromatogram, so *im*-DNP-His not observed.

contained both an acidic and a basic residue. Thus [Glu.Arg] (*B*1 γ 3) and Glu.His (*B*1 γ 1) were adsorbed, whereas Arg.Gly (*B*2 γ 5) and Ser.[His, Leu]

Structure of the peptides

The most probable composition of fraction *B* is given in Table 14. This is obtained from a consideration of the composition of insulin (Tristram, 1949) and of the very approximate analysis of fraction *A* (Sanger, 1949*a*), and may therefore not be very accurate. The value for cysteic acid was determined directly on fraction *B* by Rees by determination of the S content, and the figures for histidine, arginine, lysine and threonine are reliable since fraction *A* contains none of these residues.

N-Terminal peptide. The results with the DNP peptides demonstrated the presence of the *N*-terminal sequence, Phe.Val.Asp.Glu (Sanger, 1949*b*). The present results offer ample confirmation for this sequence, which may be deduced as follows:

Table 14 indicates that fraction *B* contains only one aspartic acid residue, so that all peptides con-

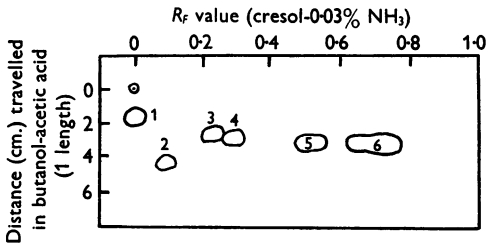


Fig. 16. Chromatogram of fraction *B*5 γ (see Table 13).

(*B*2 γ 7) were not. Similarly, it was found that the DNP derivative of Glu.Arg could be adsorbed from acid solution on to talc, whereas the derivative of

Table 13. Peptides from fraction *B*5 γ

Spot no. (Fig. 16)	Amino-acids present	Strength of amino-acid after		Structure
		Hydrolysis	DNP treatment and hydrolysis	
1	Glu	x x x	x x x	Gly.[Glu, Orn]
	Gly	x x x	x	
	Orn	x x x x	x x x *	
2	Gly	x		Glycine
3	Orn	x		Ornithine
4	Lys	x		Lysine
5	Ala	x x		Lys.Ala (<i>B</i> 2 γ 4)
	Lys	x x		
6	Ala	x x x	x x x	Pro.[Ala, Lys]
	Lys	x x x	x x *	
	Pro	x x x	?	

* DNP derivative.

Spot 1. DNP derivative hydrolysed 6 hr. This peptide was also found in the neutral fraction (*B*5 β), which is where one would expect to find it.

Spot 6. DNP derivative hydrolysed 12 hr. with 11*N*-HCl at 105°.

Arg.Gly was only weakly adsorbed, in contrast to the behaviour of most DNP peptides.

Probably the best method of carrying out the preliminary group separations on these hydrolysates would be to obtain first the acidic, neutral and basic fractions by ionophoresis. The acidic one could be further fractionated first with charcoal and then on the ion-exchange resin to separate peptides of cysteic acid. The neutral one could be separated into two fractions by charcoal adsorption, and the basic one could be subjected to a second ionophoresis at a higher pH to separate arginine peptides.

From the paper chromatograms it may be concluded that in general mixtures containing up to about twenty peptides (fractions *B*2 α and *B*2 β) may be satisfactorily fractionated, though for preliminary work it is preferable to have fewer present.

Table 14. Probable amino-acid composition of fraction *B*

Amino-acid	No. of residues/mol.
Cysteic acid	2
Aspartic acid	1
Glutamic acid	3-4
Serine	1
Glycine	3
Threonine	1
Alanine	2
Tyrosine	2
Valine	2
Leucine	3-4
Phenylalanine	3
Proline	1
Histidine	2
Lysine	1
Arginine	1

taining aspartic acid should fit into a single sequence. Six such peptides have been found: [Asp, Glu] (*B*1 α 2), Val.Asp (*B*1 β 8), Val.[Asp, Glu] (*B*1 β 10), Phe.[Val, Asp, Glu] (*B*1 β 12), Phe.[Val, Asp] (*B*1 β 13) and Phe.[Asp, Glu, His, Val] (*B*1 γ 6). It can be seen that the structure of the first five peptides can only be explained on the basis of the above sequence. Phe.Val (*B*4 β 2) is also derived from this sequence.

Peptide *B*1 γ 6 moves at different rates on the chromatograms from the other peptides of aspartic acid so must contain histidine as an integral part and not as an impurity. It can also be seen that none of the other histidine peptides would be in this position. Its structure must therefore be Phe.Val.-Asp.Glu.His. This is also supported by the identification of Glu.His (*B*1 γ 1). The fact that no peptides of aspartic acid have been found that do not fit this sequence confirms that fraction *B* contains only one aspartic acid residue.

Peptides of lysine. With the DNP method it was shown that the lysine was present in the sequence Thr.Pro.Lys.Ala, and this sequence is also confirmed in the present work by the identification of the following peptides: Thr.Pro (*B*1 δ 8), Lys.Ala (*B*2 γ 4), Thr.[Pro, Lys, Ala] (*B*2 γ 6), Pro.[Lys, Ala] (*B*5 γ 6). Since fraction *B* only contains one lysine and one threonine residue this establishes the sequence. No other peptides containing proline have been found, indicating that there is also only one proline residue. It appears that the Thr.Pro bond is unusually resistant to concentrated hydrochloric acid since large amounts of Thr.Pro and the tetrapeptide were found, whereas only traces of free proline could be detected. The tripeptide Pro.Lys.-Ala was not found in the concentrated acid hydrolysate, whereas it was in the alkaline hydrolysate.

Peptides of arginine. Two dipeptides of arginine were detected: Glu.Arg (*B*1 γ 3) and Arg.Gly (*B*2 γ 5). Since there is only one arginine residue this establishes the sequence Glu.Arg.Gly. In peptide *B*5 γ 1 (Gly.[Glu, Orn]) the ornithine is evidently derived from arginine. Since glycine is the *N*-terminal residue its structure must be Gly.Glu.Orn, and the order in fraction *B* is therefore Gly.Glu.-Arg.Gly. Gly.Glu (*B*1 δ 1) was detected also and appears to originate from this sequence. This structure could give rise to two tripeptides containing glycine, glutamic acid and arginine. Peptide *B*1 γ 2 is either one or both of these. Difficulties in deaminating this peptide may have been due to its being a mixture.

Peptides of cysteic acid. Three dipeptides containing cysteic acid have been detected: CySO₃H.Gly (*B*1 α 1), Val.CySO₃H (*B*1 α 4) and Leu.CySO₃H (*B*1 α 6). Since fraction *B* contains only two cysteic acid residues, all peptides in which cysteic acid is not

N-terminal must therefore contain the sequences Val.CySO₃H or Leu.CySO₃H. Peptide *B*1 α 3 is therefore Val.CySO₃H.Gly and peptide *B*1 α 5 is Leu.CySO₃H.Gly. The only possible structure for *B*1 α 8, which contains valine but no glycine is Leu.-Val.CySO₃H and this was also demonstrated by partial hydrolysis. This establishes the presence of the sequence Leu.Val.CySO₃H.Gly in fraction *B*, and peptide *B*1 α 7 clearly has this structure. Leu.-Val (*B*3 β 7) has also been detected. Peptide *B*1 γ 4 which contains no valine or glycine can only be His.Leu.CySO₃H. This establishes a second tetrapeptide sequence His.Leu.CySO₃H.Gly, and peptide *B*4 β 1 presumably has this structure. The presence of His.Leu (*B*2 γ 8) is further evidence for this sequence. Another peptide containing cysteic acid is *B*1 β 5 ([CySO₃H, Glu, His, Leu]) and the question arises as to whether it is a pure peptide or could be a mixture. The cysteic acid cannot be present as the free amino acid, which moves slower on phenol; therefore it must be in a peptide. No valine or glycine is present so that it must be present in the sequence Leu.CySO₃H. The dipeptide Leu.CySO₃H goes considerably faster on butanol-acetic acid and would appear below spot 6 (Fig. 6), so that the cysteic acid in spot 5 must be present in the form His.Leu.-CySO₃H. This tripeptide (*B*1 γ 4) goes faster on phenol (Fig. 7) so that the cysteic acid must be present in the tetrapeptide Glu.His.Leu.CySO₃H. In support of this structure we have also found the peptides Glu.His (*B*1 γ 1), and [Glu, His, Leu] (*B*1 γ 7). The sequence in fraction *B* is therefore Glu.His.Leu.CySO₃H.Gly. The fact that spot *B*2 α 8 ([CySO₃H, Glu, His, Leu]) sometimes contains a trace of glycine suggests that the pentapeptide may also be present.

Peptide *B*2 α 21 ([CySO₃H, Tyr, Val, Leu]) moves considerably faster on cresol than any of the other cysteic acid peptides including Leu.Val.CySO₃H (*B*2 α 17) (Fig. 11). The tyrosine must therefore be present as part of the peptide, and since it contains no glycine or histidine the cysteic acid must be in the sequence Leu.Val.CySO₃H and must be the *C*-terminal residue. The structure of the peptide is thus Tyr.Leu.Val.CySO₃H, and the order in fraction *B* must be Tyr.Leu.Val.CySO₃H.Gly. This sequence would be expected to give rise to Tyr.Leu and Tyr.Leu.Val. Spots *B*3 β 10 ([Tyr, Leu]) and *B*3 β 12 ([Tyr, Val, Leu]) probably contain these peptides, though it is not certain that they are not mixtures.

Peptides of valine and alanine. According to Table 14 fraction *B* should only contain two valine residues. One of these is present in the *N*-terminal sequence Phe.Val.Asp.Glu.His, another in the sequence Tyr.Leu.Val.CySO₃H.Gly. However there are other peptides containing valine which do not fit these sequences. It is therefore necessary to assume either that fraction *B* contains three valine

Table 15. *Peptides obtained from phenylalanyl chain of insulin (fraction B)*

Dipeptides	Phe. Val. Asp. Glu. His. Leu. CySO ₃ H. Gly (B 4β2) (B 1α2) (B 2γ8) (B 1α1)	Thr. Pro. Lys. Ala (B 188) (B 2γ4)	Gly. Glu. Arg. Gly (B 181) (B 2γ5)	Tyr. Leu. Val. CySO ₃ H (B 3β10) (B 1α4)	Ser. His. Leu. Val. Glu. Ala (B 2γ2) (B 3β7) (B 184)	Ala. Leu (B 1811)
	Val. Asp. Glu. His. Leu. CySO ₃ H (B 188) (B 1γ1) (B 1α6)		Glu. Arg (B 1γ3)	Leu. Val. CySO ₃ H. Gly (B 3β7) (B 1α1)	His. Leu. Val. Glu (B 2γ8) (B 187)	Gly. Phe (B 1γ12)
Tripeptides	Phe. Val. Asp (B 1β13)	His. Leu. CySO ₃ H (B 1γ4)	Pro. Lys. Ala (B 5γ6)	Tyr. Leu. Val (B 3β12)	Ser. His. Leu (B 2γ7)	Ala. [Tyr. Leu] (B 3β8)
	Val. Asp. Glu (B 1β10)	Leu. CySO ₃ H. Gly (B 1α5)	[Glu. Arg. Gly] (B 1γ2)	Leu. Val. CySO ₃ H (B 1α8)	Val. Glu. Ala (B 1γ10)	
		Glu. His. Leu (B 1γ7)		Val. CySO ₃ H. Gly (B 1α3)	Leu. Val. Glu (B 2β14)	
Higher peptides	Phe. Val. Asp. Glu (B 1β12)	Thr. Pro. Lys. Ala (B 2γ6)		Tyr. Leu. Val. CySO ₃ H (B 2α21)	Ser. His. Leu. Val (B 4γ2)	
				Leu. Val. CySO ₃ H. Gly (B 1α7)	Leu. Val. Glu. Ala (B 2β15)	
					His. Leu. Val. Glu (B 2β18)	
	Phe. Val. Asp. Glu. His (B 1γ6)				Ser. His. Leu. Val. Glu (B 4γ3)	
	Glu. His. Leu. CySO ₃ H (B 1β5)					
Sequence in fraction B	Phe. Val. Asp. Glu. His. Leu. CySO ₃ H. Gly	Thr. Pro. Lys. Ala	Gly. Glu. Arg. Gly	Tyr. Leu. Val. CySO ₃ H. Gly	Ser. His. Leu. Val. Glu. Ala (B 4γ4)	Others

residues or else that it is not homogeneous. The former explanation is preferred.

Besides the sequences already considered, valine is present in the following peptides: Val. Glu (*B*187), Val.[Ala, Glu] (*B*1710), Leu.[Glu, Val] (*B*2814) and Leu.[Val, Glu, Ala] (*B*2815). The only sequence fitting all these peptides is Leu.Val.Glu.Ala. This sequence may also be deduced from a consideration of the peptides of alanine. Fraction *B* contains two alanine residues; one is at the end of the sequence Thr.Pro.Lys.Ala. Other peptides are Ala.Leu (*B*1811), Glu.Ala (*B*184), Val.[Ala, Glu] (*B*1710), Leu.[Val, Glu, Ala] (*B*2815) and Ala.[Tyr, Leu] (*B*389). Any peptide containing alanine, which is not *N*-terminal, must either contain Lys.Ala or Glu.Ala. *B*1710 is therefore Val.Glu.Ala and *B*2815 is Leu.Val.Glu.Ala. The position of Ala.Leu or of Ala.[Tyr, Leu] cannot be deduced and they need not necessarily be from the same sequence.

Peptides of histidine. Two sequences containing histidine have already been considered: Phe.Val.-Asp.Glu.His and Glu.His.Leu.CySO₃H.Gly. Besides Glu.His and His.Leu, the only other dipeptide of histidine is Ser.His (*B*272). Since there are only two histidine residues and one serine residue in fraction *B*, the tripeptide *B*277 (Ser.[His, Leu]) can only be Ser.His.Leu. The above two pentapeptide sequences, which both contain Glu.His but not Ser.His must therefore be parts of the same sequence, so that the actual order in fraction *B* is Phe.Val.Asp.Glu.His.Leu.CySO₃H.Gly. Further evidence for this sequence was obtained by an extension of the DNP technique. A dilute-acid hydrolysate of the DNP derivative of fraction *B* was fractionated on a buffered Hyflo column using ethyl acetate. Two slow-moving bands were present; one contained DNP-phenylalanine, valine, aspartic acid, glutamic acid, *im*-DNP-histidine and leucine and the second contained these six residues together with cysteic acid and glycine. This in itself could not be considered as direct evidence as they could have been mixtures of terminal DNP-phenylalanyl peptides with peptides of *im*-DNP-histidine, but it adds support to the above conclusion. The establishment of this sequence depends chiefly on the two large peptides *B*176 (Phe.[Val, Asp, Glu, His]) and *B*185 ([Glu, His, Leu, CySO₃H]) which are present in rather small amounts and therefore cannot be regarded as completely proved. Ample confirmation was, however, obtained from the results with enzymic hydrolysates (Sanger & Tuppy, 1951).

Further peptides containing histidine were present in the dilute-acid hydrolysate (fraction *B*47). Peptide *B*472 (Ser.[His, Val, Leu]) must contain the sequence Ser.His.Leu and is therefore Ser.His.-Leu.Val. The four peptides of fraction *B*47, shown in Fig. 15, gave a yellow-grey colour with ninhydrin which is characteristic of peptides containing serine

as the *N*-terminal residue, and the DNP results showed this to be the case for *B*474. Thus the structure of *B*473 must be Ser.His.Leu.Val.Glu and of *B*474 Ser.His.Leu.Val.Glu.Ala. This sequence may also be derived by considering the valine peptides. The three valine residues are present as Phe.Val.Asp, Tyr.Leu.Val.CySO₃H, Leu.Val.Glu.Ala. The latter is the only one to which Ser.His.Leu.Val could be combined, so that the order in fraction *B* must be Ser.His.Leu.-Val.Glu.Ala. Spot *B*2818 ([Glu, His, Val, Leu]) is probably His.Leu.Val.Glu derived from this sequence.

Other peptides. In Table 15 are listed all the peptides detected in this work and the sequences which have been deduced in the foregoing discussion. We have included all peptides found in the hydrolysates with the exception of some of the fast-moving peptides such as are present in spots *B*2824 and *B*2825, which could not be identified. Apart from these sequences there are only three peptides which are not included in the five sequences listed. These are Ala.Leu, Gly.Phe, and Ala.[Tyr, Leu].

According to Table 14 all the amino-acids of fraction *B* can be accounted for in the above sequences except for one tyrosine, two phenylalanine and possibly one glutamic acid and one leucine residue.

From Table 15 it can be seen that of the various dipeptides that would be produced on breakdown of the five sequences, only one (Pro.Lys) was not detected. These results are in complete accord with the view that there is only one polypeptide chain in fraction *B*. All peptide spots present in detectable amounts have been taken into account and no peptides have been found that contradict this view. This would be almost impossible if there were two different chains in fraction *B* or if there was a significant amount of impurity.

SUMMARY

1. Partial hydrolysates of fraction *B* of oxidized insulin (Sanger, 1949*a*) have been fractionated by paper chromatography, and the structure of the resulting peptides determined.

2. It is concluded that the following amino-acid sequences are present in this fraction (abbreviations are given in Table 1, p. 463): Phe.Val.Asp.Glu.-His.Leu.CySO₃H.Gly, Thr.Pro.Lys.Ala, Gly.-Glu.Arg.Gly, Tyr.Leu.Val.CySO₃H.Gly and Ser.-His.Leu.Val.Glu.Ala.

We wish to express our thanks to the Medical Research Council for a grant to cover the cost of the insulin and to Dr R. L. M. Synge for giving us details of the four-compartment ionophoresis cell prior to publication. One of us (H.T.) wishes to thank the British Council for a Scholarship.

REFERENCES

- Brand, E. & Edsall, J. T. (1947). *Ann. Rev. Biochem.* **16**, 224.
 Consden, R., Gordon, A. H. & Martin, A. J. P. (1944). *Biochem. J.* **38**, 224.
 Consden, R., Gordon, A. H. & Martin, A. J. P. (1947). *Biochem. J.* **41**, 590.
 Consden, R., Gordon, A. H. & Martin, A. J. P. (1948). *Biochem. J.* **42**, 443.
 Consden, R., Gordon, A. H. & Martin, A. J. P. (1949). *Biochem. J.* **44**, 548.
 Consden, R., Gordon, A. H., Martin, A. J. P. & Synge, R. L. M. (1947). *Biochem. J.* **41**, 596.
 Dent, C. E. (1947). *Biochem. J.* **41**, 240.
 Dent, C. E. (1948). *Biochem. J.* **43**, 169.
 Desnuelle, P. & Casal, A. (1948). *Biochim. Biophys. Acta*, **2**, 64.
 Jones, T. S. G. (1949). *Discuss. Faraday Soc.* **7**, 285.
 Macpherson, H. T. (1946). *Biochem. J.* **40**, 470.
 Partridge, S. M. (1948). *Biochem. J.* **42**, 238.
 Partridge, S. M. & Davis, H. F. (1950). *Nature, Lond.*, **165**, 62.
 Phillips, D. M. P. (1948). *Nature, Lond.*, **161**, 53.
 Sanger, F. (1949a). *Biochem. J.* **44**, 126.
 Sanger, F. (1949b). *Biochem. J.* **45**, 563.
 Sanger, F. (1949c). *Cold Spr. Harb. Sym. quant. Biol.* **14**, 153.
 Sanger, F. & Tuppy, H. (1951). *Biochem. J.* **49**, 481.
 Synge, R. L. M. (1943). *Chem. Rev.* **32**, 135.
 Synge, R. L. M. (1951). *Biochem. J.* (in the Press).
 Theorell, H. & Åkeson, A. (1942). *Ark. Kemi Min. Geol.* **16A**, no. 8.
 Tiselius, A., Drake, B. & Hagdahl, L. (1947). *Experientia*, **3**, 651.
 Tristram, G. R. (1949). *Advanc. prot. Chem.* **5**, 83.
 Wieland, T. & Wirth, L. (1949). *Ber. dtsch. chem. Ges.* **82**, 468.

The Amino-acid Sequence in the Phenylalanyl Chain of Insulin

2. THE INVESTIGATION OF PEPTIDES FROM ENZYMIC HYDROLYSATES

BY F. SANGER (Beit Memorial Fellow) AND H. TUPPY*

Biochemical Laboratory, University of Cambridge

(Received 17 January 1951)

In the previous paper (Sanger & Tuppy, 1951) a large number of simple peptides in partial hydrolysates of fraction B of oxidized insulin were identified and several amino-acid sequences were deduced from them. It was not possible to determine the complete structure of the polypeptide chain, partly on account of the difficulty of fractionating the less polar peptides containing the aromatic amino-acids and partly because of the relatively great lability of certain bonds such as those involving the amino groups of the hydroxyamino-acids.

In the present paper are described investigations of the peptides produced on hydrolysis of fraction B with the proteolytic enzymes, pepsin, chymotrypsin and trypsin. These peptides, which are mostly larger than those present in the acid and alkaline hydrolysates, could be successfully fractionated by paper chromatography, and with the knowledge presented in the previous paper it was possible to determine the structure of most of them and to work out the complete sequence of fraction B and hence of the phenylalanyl chain of insulin.

The abbreviations, etc., defined in the previous paper (Sanger & Tuppy, 1951) will be used throughout and the peptides with reference numbers starting B 1, B 2, B 3, B 4, B 5 refer to those described in that paper.

* Present address: II. Chemisches Universitätslaboratorium, Vienna, Austria.

METHODS

Enzymes

Crystalline pig pepsin was obtained from Armour Laboratories. The chymotrypsin was prepared and several times recrystallized by Mr J. Lightbown of the National Institute for Medical Research, Mill Hill, London. The trypsin was a crystalline preparation from Armour Laboratories. It was recrystallized three times according to the method of Kunitz & Northrop (1936).

Hydrolyses

Experiment Bp (pepsin). Fraction B (30 mg.) was dissolved in 3 ml. 0.01 M-HCl. A solution of 0.3 mg. pepsin in 0.3 ml. 0.01 M-HCl was added, and the mixture incubated at 37° for 24 hr. After boiling to inactivate the pepsin, the solution was taken to dryness *in vacuo* to remove excess HCl and subjected to ionophoresis in the four-compartment cell of Synge described in the previous paper. This separated the hydrolysate into acid, neutral and basic fractions.

Experiment Bc (chymotrypsin). Fraction B (50 mg.) was dissolved in 5 ml. water and brought to pH 7.5–8.0 by the addition of NaHCO₃. To this was added 2.3 mg. chymotrypsin, the pH again adjusted and the solution incubated at 25° for 24 hr. It was then brought to about pH 3 by addition of HCl and taken to dryness in a desiccator. The hydrolysate was again separated into acidic, neutral and basic fractions by ionophoresis.

Experiment Bt (trypsin). Fraction B (25 mg.) was dissolved in 2.5 ml. water and the pH adjusted to 7.5 with NaHCO₃. Trypsin (1.2 mg.) dissolved in 1.2 ml. water was added and the mixture incubated at 37° for 24 hr. On acidification a